

Evaluation of the performance of Euro-CORDEX regional climate models for assessing hydrological climate change impacts in Great Britain

Pastén-zapata, Ernesto; Jones, Julie M.; Moggridge, Helen; Widmann, Martin

DOI:

[10.1016/j.jhydrol.2020.124653](https://doi.org/10.1016/j.jhydrol.2020.124653)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Pastén-zapata, E, Jones, JM, Moggridge, H & Widmann, M 2020, 'Evaluation of the performance of Euro-CORDEX regional climate models for assessing hydrological climate change impacts in Great Britain: a comparison of different spatial resolutions and quantile mapping bias correction methods', *Journal of Hydrology*, vol. 584, 124653, pp. 1-17. <https://doi.org/10.1016/j.jhydrol.2020.124653>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Journal Pre-proofs

Research papers

Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: a comparison of different spatial resolutions and quantile mapping bias correction methods

Ernesto Pastén-Zapata, Julie M. Jones, Helen Moggridge, Martin Widmann

PII: S0022-1694(20)30113-X
DOI: <https://doi.org/10.1016/j.jhydrol.2020.124653>
Reference: HYDROL 124653

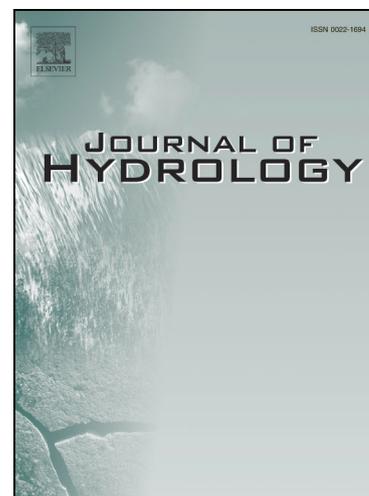
To appear in: *Journal of Hydrology*

Received Date: 9 July 2019
Revised Date: 16 January 2020
Accepted Date: 1 February 2020

Please cite this article as: Pastén-Zapata, E., Jones, J.M., Moggridge, H., Widmann, M., Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: a comparison of different spatial resolutions and quantile mapping bias correction methods, *Journal of Hydrology* (2020), doi: <https://doi.org/10.1016/j.jhydrol.2020.124653>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.



1 Title:

2 Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological
3 climate change impacts in Great Britain: a comparison of different spatial resolutions and quantile
4 mapping bias correction methods

5 Authors:

6 Ernesto Pastén-Zapata^{1,2*}, Julie M. Jones², Helen Moggridge², Martin Widmann³

7 Affiliations:

8 ¹ Geological Survey of Denmark and Greenland (GEUS), Department of Hydrology, Denmark

9 ² Department of Geography, University of Sheffield, United Kingdom

10 ³ School of Geography, Earth and Environmental Sciences, University of Birmingham, United Kingdom

11

12 * Corresponding author:

13 Ernesto Pastén-Zapata
14 Department of Hydrology
15 Geological Survey of Denmark and Greenland (GEUS)
16 Oester Voldgade 10
17 1350 Copenhagen K
18 Denmark
19 epz@geus.dk
20

21

22 Regional Climate Models (RCMs) are an essential tool for analysing regional climate change impacts,
23 such as hydrological change, as they provide simulations with more small-scale details and expected
24 smaller errors than global climate models. There has been much effort to increase the spatial resolution
25 and simulation skill of RCMs (i.e. through bias correction), yet the extent to which this improves the
26 projection of hydrological change is unclear. Here, we evaluate the skill of five reanalysis-driven Euro-
27 CORDEX RCMs in simulating precipitation and temperature, and as drivers of a hydrological model to
28 simulate river flow on four UK catchments covering different physical, climatic and hydrological
29 characteristics. We use a comprehensive range of evaluation indices for aspects of the distribution such
30 as means and extremes, as well as for the structure of time series. We test whether high-resolution RCMs
31 provide added value, through analysis of two RCM resolutions, 0.44° (50 km) and 0.11° (12.5 km), which
32 are also bias-corrected employing the parametric quantile-mapping (QM) method, using the normal
33 distribution for temperature, and the Gamma (GQM) and Double Gamma (DGQM) distributions for
34 precipitation. The performance of these is considered for a range of meteorological variables and for the
35 skill in simulating hydrological impacts at the catchment scale.

36 In a small catchment with complex topography, the 0.11° RCMs clearly outperform their 0.44° version
37 for precipitation and temperature, but when used in combination with the hydrological model, fail to
38 capture the observed river flow distribution. In the other (larger) catchments, only one high-resolution
39 RCM consistently outperforms its low-resolution version, implying that in general there is no added value
40 from using the high-resolution RCMs in those catchments. Both resolutions produce river flow simulations
41 that cover the observed flow duration curve, but the ensemble spread is large and therefore the
42 simulations are difficult to use in practice. GQM decreases most of the simulation biases, except for
43 extreme precipitation and high flows, which are further decreased by DGQM, which also reduces the
44 multi-model simulation spread. Bias correction does not improve the representation of daily temporal
45 variability measured by the Nash-Sutcliffe Efficiency Index, but it does for monthly variability, in particular
46 when applying DGQM, which reduces most of the simulation biases. Overall, an increase in RCM resolution
47 does not imply a better simulation of hydrology and bias-correction represents an alternative to ease
48 decision-making.

49

50

51

52 1. Introduction

53 Global General Circulation Models (GCMs) are the main tool for climate change projections. However,
54 their spatial resolution is usually not finer than 100 km x 100 km (Rummukainen, 2016), limiting their skill
55 to simulate local climate. Regional Climate Models (RCMs) focus on specific subcontinental or subnational
56 domains, incorporating regional features such as topography, coasts and islands more accurately.
57 Consequently, RCMs improve the simulation of small-scale processes that affect precipitation, such as
58 orographic forcing (Rummukainen et al., 2015; Di Luca et al., 2015), and are expected to yield more
59 accurate projections of climate change at finer spatial scales. RCMs have been used extensively to evaluate
60 the impacts of climate change on hydrology, such as changes in mean river flow, floods or low flows (e.g.
61 Kay et al., 2015; Kay and Jones, 2012; Mendoza et al., 2016; Teng et al., 2015; Prudhomme et al., 2013;
62 Cloke et al., 2013).

63 The resolution of RCMs has increased over time with the availability of higher computer power.
64 Currently, the spatial resolution of RCMs varies from 50 km x 50 km to less than 5 km x 5km (Rummukainen,

65 2016; Rockel et al., 2015). Due to their increased representation of regional features and small-scale
66 processes, RCMs generally simulate the current regional climate better than their driving data (Feser et al.
67 2011; Di Luca et al., 2015). Nevertheless, this might not be true in regions mainly influenced by large-scale
68 climatic processes (Eden et al., 2014). Therefore, the added value of high-resolution RCMs depends on
69 the analysed region, variable and context (Rummukainen, 2016).

70 An important driver for increasing RCM resolution is the need to improve the analysis of climate change
71 impacts for decision-making (e.g. Macadam et al., 2016; Qian et al., 2015). For hydrology, the standard
72 analysis of climate change impacts generally involves coupling uncorrected or bias-corrected GCM or RCM
73 precipitation and temperature outputs with hydrological models to simulate river flow scenarios (e.g.
74 Teutschbein and Seibert; 2012; Huang et al., 2014; Teng et al., 2015). In Great Britain, these studies focus
75 on one (or more) of four main topics: 1) the contribution of the GCMs, RCMs, emission scenarios and bias-
76 correction techniques to the uncertainty of the change projection (e.g. Prudhomme and Davies, 2009; Kay
77 et al., 2009; Arnell, 2011; Christerson et al., 2012), 2) the impact of the bias correction techniques on the
78 projections (e.g. Prudhomme et al., 2013; Cloke et al., 2013; Wetterhall et al., 2012; Kim et al., 2016), 3)
79 projections of future floods (Cloke et al., 2013; Kay et al., 2015; Wetterhall et al., 2012; Kay and Jones,
80 2012), and, 4) projections of future low flows (Wilby and Harris, 2006; Arnell, 2011; Fowler and Kilsby,
81 2007).

82 Some studies have identified a consistent improvement in hydrological simulation skill with increasing
83 RCM resolution for the annual mean river flow (Huang et al. 2014). For the simulation of river flow peaks
84 as a response to extreme precipitation events, previous studies found no improvement when increasing the
85 model resolution (Kay et al. 2015; Huang et al., 2014). Others studies found that the improvement depends
86 on the catchment size and on the evaluation index (Dankers et al. 2007), whilst others found an
87 improvement when simulating seasonal flow and hydrologic signatures aimed to represent diverse
88 hydrologic processes (e.g. runoff ratio, center time of runoff) (Mendoza et al., 2016). However, these studies
89 have only used one RCM to perform the comparison as, to date, there has been no systematic study using
90 a large number of RCM simulations to test the effect of RCM resolution on hydrological simulation skill.

91 The first aim of this paper is to use the EURO-CORDEX simulations (Jacob et al. 2014) to robustly
92 assess the added value of increasing RCM resolution on hydrological simulations. The Euro-CORDEX
93 simulations at 0.11° (12.5 km x 12.5 km) and 0.44° (50 km x 50km) have the same lateral boundaries and
94 the parameterisations of each RCM are the same at both resolutions, thus making them ideal for such a
95 comparison. This work builds on assessments of the 0.11° and 0.44° Euro-CORDEX RCMs at reproducing
96 observed temperature and precipitation distributions, including extremes and dry/wet spell lengths. Results
97 vary among the studies. Some found a higher accuracy for the 0.11 RCMs for Europe when evaluating the
98 mean and extreme precipitation at a daily and sub-daily temporal resolution (Prein et al. 2015, Fantini et al
99 2016), whereas others did not find an improvement in accuracy when assessing the spatio-temporal
100 patterns of the monthly and seasonal precipitation and temperature (Kotlarski et al. 2014). For the Alps
101 Torma et al. (2015) found a higher skill for the 0.11 RCMs when simulating the spatial distribution of the
102 mean, extreme and intensity of precipitation, while Casanueva et al. (2016) showed for the Alps and Spain
103 that the best performance depends on the RCM, season and validation index when evaluating precipitation
104 intensity, frequency, mean and extremes.

105 Biases in RCM simulations are due to parameterisation of sub-grid processes, limited representation
106 of local features, incorrect boundary conditions and differences between spatial resolutions of the
107 simulations and observations (Ehret et al., 2012; Benestad, 2010). Therefore, RCMs require post-
108 processing for many applications (Christensen et al., 2008). Statistical bias-correction techniques reduce
109 biases in the mean, variance or the complete distribution of simulated climate variables (reviews in Maraun
110 et al., 2010; Teutschbein and Seibert, 2012; Maraun and Widmann, 2018; Lafon et al., 2013). Quantile
111 mapping (QM) is one of the standard techniques used (Piani et al., 2010; Teutschbein and Seibert, 2012;
112 Maurer et al., 2014). Whilst effective, bias correction has important limitations that are further discussed in
113 the conclusions.

114 To date, a detailed comparison of the simulation skill of bias-corrected high- and low-resolution model
115 outputs for aspects that are important for hydrological studies (e.g. means, extremes, daily sequence) has
116 not been undertaken. The second aim of this study addresses this research gap by conducting a detailed
117 evaluation of aspects that are relevant for the hydrological regime such as seasonal precipitation,

118 occurrence of extreme events, and monthly and daily pairwise indices (assess the skill to reproduce the
119 observed time-series). The evaluation of these aspects allows identifying the capabilities and weaknesses
120 of the impact assessments. Here, the simulations are evaluated against gauged data, working as a mean
121 to assess the plausibility of the simulation outputs using uncorrected and bias-corrected RCMs. This work
122 builds on studies that have assessed climate variables. For instance, the bias-corrected Euro-CORDEX
123 simulations, at both resolutions, have a similar skill at capturing the wet-day intensity and precipitation
124 frequency (Casanueva et al., 2016).

125 Here, we therefore address the two above-mentioned research aims by evaluating the simulation skill
126 of five uncorrected and bias-corrected Euro-CORDEX RCMs at 0.11° and 0.44° using a range of
127 temperature, precipitation and river flow indices, evaluating the mean along with high and low extremes,
128 frequency of occurrence and daily and monthly simulation sequence. By using a multi-model ensemble,
129 this analysis provides a robust understanding of the added value of high-resolution simulations and post-
130 processing approaches for hydrological impact studies. We analyse four diverse catchments across Great
131 Britain, representative of different climate and physical characteristics, focusing on the following questions:

- 132 1) Based on a range of selected indices, is the performance of the 0.11° Euro-CORDEX RCMs better
133 than their 0.44° version to simulate (a) climate and (b) river flow?
- 134 2) Is the current skill of the Euro-CORDEX RCMs sufficient to generate plausible inputs for the
135 analysis of climate change impacts on hydrology and how does this compare to the inputs from
136 bias-corrected simulations?
- 137 3) Is there any improvement in the simulation skill of precipitation and river flow when using a Double
138 Gamma Quantile Mapping (DGQM) bias correction compared to the usual Gamma Quantile
139 Mapping (GQM) approach?

140 Given the associated computational cost (Bucchignani et al., 2016) and the potential for improving the
141 skill of climate simulations, especially for impact assessments (Ehret et al., 2012), there is a clear need for
142 rigorous evaluation of the added value of increasing RCM resolution. Previous hydrological impact studies
143 have analysed this issue using one or two RCMs (e.g. Mendoza et al., 2016; Kay et al., 2015). However,
144 their results might not be transferable to other RCMs, as each has its own parameterisations.

145 GQM inflates the precipitation extremes, producing unreliable flood simulations. Whilst this is a known
146 issue (Cloke et al., 2013; Huang et al., 2014), no study has exhaustively compared the results between
147 using the GQM and the DGQM approaches using extreme indices. This study provides a comprehensive
148 analysis of such gaps.

149 **2. Data and method**

150 **2.1. Observation databases and study catchments**

151 The observations are used to calibrate the hydrological model (Section 2.2), develop the bias
152 correction method (Section 2.3) and to compare the outputs of the RCMs to evaluate their simulation skill
153 (Section 2.5). We employ gridded observations based on weather stations, as these are better comparable
154 to the outputs of the climate models which produce an areal average for each gridbox (following Osborn
155 and Hulme, 1997). We use the Centre for Ecology and Hydrology (CEH) Gridded Estimates of Areal Rainfall
156 (CEH-GEAR) dataset (Tanguy et al. 2014) as 1km x 1km gridded daily precipitation observations (Keller et
157 al., 2015). Records from the Natural Environment Research Council (NERC) Hydrology and Ecology
158 Research Support System (CHESS) (Robinson et al., 2017a, 2017b) are used as 1 km x 1 km gridded daily
159 temperature observations. The 1 km x 1 km gridded CHESS-PET dataset is employed as potential
160 evapotranspiration (PET) observational reference. CHESS-PET uses the Penman-Monteith equation
161 (Monteith, 1965) to calculate daily PET using climate variables from the Met Office Rainfall and Evaporation
162 Calculation System (MORECS) (Hough and Jones, 1997) as input. All these datasets cover the period 1961
163 to 2010. A detailed description of the methodology and weather stations used to develop the gridded
164 datasets can be found in Robinson et al. (2017a, 2017b) and Tanguy et al. (2014). We use river flow
165 observations from the CEH's National River Flow Archive (NRFA). The available river flow observations for
166 the 1961-2010 period varies in each catchment, with a minimum of 30 years of continuous records.

167 We analyse four catchments within the UK. The catchments have long river flow records and cover
168 regions that are representative of the different climate and catchment types that can be found within the
169 UK. These are the Upper Thames, Glaslyn, Calder and Coquet catchments (Fig. 1). This set of catchments
170 with different characteristics (Table 1) can aid identifying key features that impact on the simulation skill of
171 the RCMs. The smallest catchment is the Glaslyn, which has the most complex topography and highest

172 rainfall. The largest catchment is the Upper Thames (1616 km²), which also has the least complex
 173 topography. The Calder and Coquet are intermediate in terms of area, elevation and precipitation. These
 174 catchments have been studied before using bias-corrected climate projections (QM, normal distribution for
 175 temperature and Gamma distribution for precipitation) from the HadRM3-PPE RCM (Prudhomme et al.,
 176 2013).

177 **2.2. RCMs**

178 We evaluate two spatial resolutions (0.11° equivalent to 12.5 km x 12.5 km and 0.44° equivalent to
 179 50 km x 50 km) of five Euro-CORDEX RCMs (Jacob et al., 2014) driven by the ERA-Interim reanalysis (Dee
 180 et al., 2011), the so-called ‘evaluation simulations’. The evaluation simulations are used as these are driven
 181 by observations and consequently simulate the internal variability in synchronicity with reality, in contrast
 182 to the historical simulations. The assessed RCMs are shown in Table 2 (refer to Table 1 in Kotlarski et al.
 183 (2014) and Table 1 in Prein et al. (2015) for a detailed RCM description). These models are selected as
 184 they have the best performance to reproduce observations in the British Isles according to Kotlarski et al.
 185 (2014). When more than one RCM cell is needed to fully cover the catchment we use the mean of the cells
 186 to represent the catchment’s climate simulations (see Fig. 1).

187 **2.3. Bias correction**

188 QM is used based on parametric representations of the simulated and observed distributions (Piani
 189 et al., 2010). For each month of the year, the Gamma distribution is fitted to the observed and simulated
 190 gridded daily precipitation and the normal distribution to the observed and simulated gridded daily
 191 temperature. RCMs generally simulate too many days with very low precipitation and not enough dry days.
 192 Therefore, in an initial step the QM method adjusts the number of simulated dry days in the RCM evaluation
 193 simulations such that they match with the number of observed dry days by including a wet day threshold
 194 and replacing all values below it with zero. After the wet-day adjustment, the distributions of the simulations
 195 and observations are matched using their cumulative distribution functions (CDF). The method is
 196 represented by the following equations:

$$197 \quad P_c(t) = F_g^{-1}(F_g(P_R(t), \alpha_R, \beta_R), \alpha_O, \beta_O) \quad (1)$$

$$198 \quad T_c(t) = F_n^{-1}(F_n(T_R(t), \mu_R, \sigma_R^2), \mu_O, \sigma_O^2) \quad (2)$$

199

200 Where $P_c(t)$ and $P_R(t)$ represent the bias-corrected and raw RCM daily precipitation, respectively.

201 Likewise, $T_c(t)$ and $T_R(t)$ stand for the bias-corrected and raw RCM daily temperature. The raw RCM CDF
 202 is symbolized with F , and F^{-1} stands for the observations inverse CDF. The 'g' and 'n' subscripts represent
 203 the Gamma and normal distributions, respectively. The precipitation shape and scale parameters are
 204 symbolised by α and β and the temperature mean and standard deviation by μ and σ , respectively. Finally,
 205 the 'R' and 'O' subscripts are used to symbolize the distribution parameters from the raw RCM and
 206 observations, respectively.

207 GQM focuses on the most frequent values (e.g. means) (Teng et al., 2015; Yang et al., 2010).
 208 Consequently, the corrected precipitation extremes tend to be inflated compared to the observations
 209 (Cannon et al., 2015). Therefore, we also bias-correct precipitation using the DGQM. The methodology is
 210 mainly the same as the GQM with the difference that the simulated precipitation distribution is divided in
 211 two segments. Each is corrected separately, generating correction parameters for each section. In our
 212 study, the distribution is divided at the 90th percentile because at this percentile the biases inflate (see
 213 section 3.2.2.1).

214 For the 0.11° RCMs, the spatial scale of the simulations and the observations are approximately the
 215 same and the method can be viewed as a pure bias correction. In contrast, the output of the 0.44° is given
 216 on a larger scale than the observations and thus the QM also includes a downscaling aspect to account for
 217 the difference in distributions on different spatial scales. We note that due to the existence of sub-grid
 218 variability QM is in principle problematic as the corrected values for all sub-grid locations would have
 219 unrealistic high correlations (Maraun, 2013). However, this limitation is not of high relevance for our study
 220 as we bias-correct the distributions for the entire catchments.

221 **2.4. Hydrological simulation**

222 The Hydrological Modeling System from the US Army Hydrologic Engineering Center (HEC-HMS)
 223 (Scharffenberg, 2013) is used to simulate the catchments' daily river flow. HEC-HMS has been successfully
 224 used before to analyse climate change impacts on water resources in other regions (e.g. Babel et al., 2014;

225 Azmat et al., 2015). An advantage of the model is the available guidance for the estimation of parameters.
 226 Here, the model is run using its continuous, lumped arrangement. Observed precipitation and PET time
 227 series are used as input for the calibration and validation of the model. Afterwards, the raw and bias-
 228 corrected RCM simulations drive the model to generate the river flow simulations.

229 Evapotranspiration controls the moisture returning from the Earth's surface to the atmosphere and
 230 therefore impacts on the river flow. PET estimates the amount of water returning to the atmosphere when
 231 enough water is present in the surface of the catchment. Climate models do not simulate PET directly, thus
 232 it is estimated indirectly with formulas using variables from the climate models as input. There is no
 233 consensus on whether temperature-based or physically-based formulas provide better results in a climate
 234 change context (Kay et al., 2013) as the data required by the physically-based formulas is uncertain in the
 235 climate model simulations compared to the input from one variable formulas (Kingston et al., 2009). This
 236 has been discussed and explored elsewhere (please refer to: Seiller and Anctil, 2016; Kingston et al., 2009;
 237 Kay and Davies, 2008; Kay et al., 2013). We estimate PET using the Oudin formula (Oudin et al., 2005) as
 238 it has given accurate results before (e.g. Oudin et al., 2005; Kay and Davies, 2008).

$$239 \begin{cases} PET (mm \text{ day}^{-1}) = \frac{R_e(T+5)}{\lambda\rho\left(\frac{T+5}{100}\right)} & \text{if } T + 5 > 0 \\ PET (mm \text{ day}^{-1}) = 0 & \text{otherwise} \end{cases} \quad (5)$$

240 The extraterrestrial solar radiation (R_e) is the solar radiation received at the top of the Earth's
 241 atmosphere which can be estimated by the latitude and day of the year. The density of water is symbolized
 242 by ρ , the latent heat flux by λ (2.45 MJ/kg) (Allen et al., 1998) and T is the daily mean temperature ($^{\circ}\text{C}$).
 243 When driven by observed temperature, the Oudin formula gave results similar to the CHES-PET dataset
 244 for 1973 to 2010 (Pasten-Zapata, 2017).

245 **2.5. Hydrological model calibration**

246 The hydrological model is calibrated and validated against the observations using a split sample
 247 test. Considering the available uninterrupted daily river flow records, for each catchment two same-length
 248 non-overlapping time periods are used: one for calibration and the other for validation. The period with
 249 available river flow observations varies for each catchment. The period with observations for each
 250 catchment is selected and divided into two equal-length, non-overlapping periods. Calibration is done for

251 the more recent period and validation for the other portion of the sample. Three indices are assessed: the
 252 low flows simulation is evaluated using the Q95 (flow equalled or exceeded 95% of the time), the high flows
 253 by the Q10 (flow equalled or exceeded 10% of the time) and the Nash-Sutcliffe Efficiency Index (NSE)
 254 which evaluates the fit of the simulated and observed river flow. The NSE ranges from 1 (perfect fit) to
 255 negative (unreliable model) (Nash and Sutcliffe, 1970). In the NSE formula, Q_t^{obs} and Q_t^{sim} stand for the
 256 observed and simulated river flow at time step t , respectively. Q^{mean} is the average of the observed river
 257 flows during the complete period.

$$258 \quad NSE = 1 - \left[\frac{\sum_{t=1}^n (Q_t^{obs} - Q_t^{sim})^2}{\sum_{t=1}^n (Q_t^{obs} - Q^{mean})^2} \right] \quad (6)$$

259 **2.6. RCM validation approach and indices**

260 Validation is important to assess the RCM simulation skill before and after bias correction (Eden et
 261 al., 2014). Here, a five-fold cross-validation approach is used: 1) the study period is divided into five same-
 262 length, non-overlapping blocks, and 2) the QM methods are trained using four blocks and the remaining
 263 block is corrected using the parameters from the training period (Maraun et al., 2015). The corrected blocks
 264 are concatenated to time series for the entire period from which the performance measures for the bias-
 265 corrected precipitation and temperature are derived.

266 A range of distribution-based and time series-based indices evaluate the skill of the raw and bias-
 267 corrected RCM outputs to simulate the observed precipitation, temperature and river flow. The indices
 268 assess biases in the means, low and high extremes, inter- and intra-annual variability and correlations for
 269 each variable (see Table 3). RCMs are then ranked based on their skill to simulate all indices relative to the
 270 skill of the other RCMs at both resolutions. As we are evaluating the outputs of 10 RCMs (5 high-resolution
 271 and 5 low-resolution), each RCM is given a value between 1 (best) and 10 (worst) based on their simulation
 272 skill. Thus, simulation skill refers to the biases present in the models compared to the available
 273 observations considering all the metrics used in this study. We use the complete time series (dry days
 274 included) to estimate the precipitation indices. Even when driven by “perfect boundary conditions”, a close
 275 similarity between the RCM simulations and observations is not expected (Kay et al., 2015) due to subgrid
 276 variability or internal variability because the boundary conditions do not fully determine the weather states

277 within the RCM. Nevertheless, we include daily and monthly pairwise indices as these are important for the
278 river flow simulation. We left out the hydrological model uncertainty source intentionally to solely evaluate
279 the effects of increasing RCM resolution. Thus, we compare the river flow simulations driven by RCM
280 outputs against the river flow simulations driven by the observed temperature and precipitation.

281 **3. Results**

282 This section begins by showing hydrological model simulation skill followed by the evaluation of the
283 simulation skill of the uncorrected RCMs for temperature, precipitation and river flow. Finally, we compare
284 the biases that remain after bias-correcting precipitation using the GQM and DGQM and their impacts on
285 the river flow simulation.

286 **3.1. Calibration and validation of the hydrological model**

287 Firstly we evaluate the hydrological model simulation skill using climate observations as input.
288 Depending on the catchment, the length of the overall evaluation period ranges from 34 to 49 years. The
289 daily NSE varies between 0.62 (Calder) and 0.78 (Glaslyn) for calibration and between 0.52 (Coquet) to
290 0.78 (Glaslyn) for validation (Table 4). These results indicate a moderate to good simulation skill overall
291 compared to the NSE values from similar studies which vary from 0.45 to 0.9 (e.g. Arnell, 2011; Walsh et
292 al., 2015; Cloke et al., 2013). The Q10 bias ranges between -6% and 11% for the calibration and between
293 -5% and 7% for the validation. Similarly, the Q95 bias ranges between -27% and -11% for the calibration
294 and between -44% and 6% for the validation. Overall, the simulation of high flows is very good and moderate
295 to very good for the low flows. More detail on the calibration and validation results can be found in the work
296 from Pasten-Zapata (2017).

297 **3.2. Evaluation of the RCM simulation skill**

298 We now assess the skill of the RCMs at simulating climate and river flow, firstly for the raw
299 simulations and then for the bias-corrected outputs. We only show robust results for the analysis of the
300 indices (e.g. if all RCMs from a particular resolution underestimate or overestimate an index). We also
301 evaluate the multi-model percentile bias for each variable and use a skill rank to enable comparison of the
302 different RCMs over the different performance indices. The ranking is only estimated for the uncorrected

303 simulations as the biases after the correction are small and similar among the RCMs. Thus, ranking the
304 bias-corrected simulations would give meaningless results.

305 **3.2.1. Uncorrected RCM simulations**

306 **3.2.1.1. Temperature**

307 We begin with assessing the ability of the RCMs to simulate temperature. The 0.11° RCMs
308 underestimate the annual mean temperature for the upper Thames (Fig. 2a, ii), Calder (Fig. 2c, ii) and the
309 Coquet (Fig. 2d, ii) catchments, whereas the 0.44° RCMs overestimate the annual mean temperature for
310 the Glaslyn (Fig. 2b, ii) and Coquet (Fig. 2d, ii) catchments. The monthly mean temperature bias for the
311 0.11° RCMs is larger for the Calder (between and 0.5 °C and 1.1 °C) (Fig. 2c, ii) and smaller for the Glaslyn
312 catchment (between 0.4 °C and 0.7 °C) (Fig. 2b, ii). In contrast, the monthly mean temperature bias of the
313 0.44° RCMs is larger for the Glaslyn (between 0.4 °C and 1.2 °C) (Fig. 2b, ii) and smaller for the Calder
314 catchment (between 0.8 °C and 1.0 °C) (Fig. 2c, ii).

315 We use the simulation spread to evaluate the simulation skill of each resolution. The spread
316 represents the range between the highest and lowest simulated value considering all RCMs at each
317 resolution and all gridcells within a catchment. The temperature percentile bias spread for the upper
318 Thames is similar for both resolutions except between the 40th and 60th percentile where the 0.44°
319 simulation include larger positive biases (Fig. 3a). For the Glaslyn catchment, the 0.44° simulations
320 overestimate temperature for almost all percentiles, while the biases of the 0.11° simulations are smaller
321 (Fig. 3b). For the Calder catchment, the 0.44° RCM spread includes the no bias threshold for all percentiles,
322 whereas the 0.11° RCMs underestimate temperature between the 40th and 90th percentile (Fig. 3c). Finally,
323 in the Coquet catchment the 0.44° simulations overestimate temperature below the 70th percentile and the
324 0.11° simulations underestimate it between the 40th and 80th percentiles (Fig. 3d). The Pearson correlation
325 coefficients of the daily time series vary between 0.91 and 0.97 in all catchments for both resolutions (Figs.
326 2, iii).

327 Integrating the RCM simulation skill of all the indices into a ranking shows that, in the upper Thames,
328 two out of five high-resolution uncorrected simulations outperform their 0.44° version (last column of Table
329 5). Similarly, for the Calder catchment, one 0.11° simulation outperforms its 0.44° version and all five high-

330 resolution simulations outperform their low-resolution version for the Glaslyn and Coquet catchments. This
331 indicates that topography has an influence in the simulation of temperature and RCM resolution has an
332 effect in the simulation skill for catchments with larger elevation variability where, for observations at high
333 elevation, the 0.44° RCMs would be expected to have positive biases as the grid elevation is lower than the
334 observations.

335 Based on the rank, the overall best performing simulation for the upper Thames and Calder
336 catchments is 0.44° RACMO, whereas for the Glaslyn and Coquet catchments, the 0.11° RACMO and
337 HIRHAM simulations, respectively, outperform the rest. This implies that biases from the high-resolution
338 simulations are smaller for the catchments with complex topography, which is better represented by the
339 0.11° simulations. The biases are a consequence of systematic model biases in the elevation and a lack of
340 representation of the elevation variability. Nevertheless, for larger and flatter catchments the simulation skill
341 from both resolutions is similar.

342 **3.2.1.2. Precipitation**

343 Now we assess the skill of the uncorrected RCMs to simulate precipitation. Overall, RCMs have
344 biases when simulating extremes. For instance, the SDII ratio is underestimated in all catchments by the
345 0.44° simulations (Figs. 4a, S1a and S2a), except for the Coquet (Fig. S3a). In all catchments the RX1day
346 is overestimated by both resolutions between 24% and 93%. The R10 and R20 are underestimated at the
347 Glaslyn catchment between -23 and -77 days and between -16 and -45 days, respectively (Fig. S1d).
348 Similarly, in the Calder catchment R10 and R20 are underestimated by the 0.44° simulations between -5
349 and -10 days and between -3 and -4 days, respectively (Fig. S2d). These results indicate that the
350 uncorrected models can provide unrealistic simulations of extreme precipitation.

351 It is expected that the models simulate the precipitation mean better than the extremes. Even
352 though the spread of the models includes the observed mean precipitation for most catchments, there are
353 cases when this does not happen. The annual mean precipitation is underestimated by both resolutions in
354 the Glaslyn catchment between -22% and -67% (Fig. S1c). This may be because the analysed RCMs do
355 not correctly simulate convective precipitation. In the Calder catchment the 0.44° simulations underestimate
356 the annual mean precipitation between -7% and -16% (Fig. S2c). This can be due to local precipitation not

357 being correctly simulated by the coarse models. The absolute monthly mean precipitation bias for both
358 resolutions varies between 7% and 67% in all study cases (Figs. 4c, S1c, S2c and S3c).

359 The simulated precipitation bias spread increases in all catchments as the percentile increases.
360 The spread of the 0.11° simulations is larger than for the 0.44° simulations (Fig. 5, first row). In the upper
361 Thames catchment, the 0.11° simulations reach their largest spread, -1 to 4 mm/day, above the 95th
362 percentile whereas the largest spread of the 0.44° RCMs ranges from -1 to 1 mm/day (Fig. 5a). In the
363 Glaslyn catchment, the bias spread deviates from the observations at the 50th percentile for the 0.44°
364 simulations and at the 60th percentile for the 0.11° simulations (Fig. 5d). In the Calder catchment, the 0.11°
365 simulations spread includes the no bias threshold for the whole distribution whereas the 0.44° simulations
366 spread deviates from that threshold at the 70th percentile (Fig. 5g). In the Coquet catchment, the spread
367 from both resolutions includes the zero bias threshold for almost all percentiles (Fig. 5j).

368 The dry and wet spell biases are important for the simulation of river flow as this is influenced by
369 the daily sequence of the wet/dry conditions. The absolute dry spell bias for both resolutions in all
370 catchments range between 0.2 to 1.6 days, with a similar simulation skill in all catchments (Figs. 4b, S1b,
371 S2b, S3b). Likewise, the absolute wet spell bias for both resolutions varies between 0.1 and 1.6 days in all
372 catchments (Figs. 4b, S1b, S2b, S3b). Biases in the upper Thames for this measure are smaller, 0.2 to 0.6
373 days (Fig. 4b), compared to the other catchments. These results do not show large simulation biases.
374 Considering the time-series based indices, correlation coefficients are above 0.4 and below 0.8 in all
375 catchments, showing differences between the daily observations and simulations (Figs. 4c, S1c, S2c, S3c).

376 Considering the ranking for all indices, only for the Glaslyn catchment do all the 0.11° simulations
377 outperform their 0.44° version (Table 6, last column). From the five RCMs, two 0.11° simulations outperform
378 their low-resolution version for the Upper Thames and three for the Calder and Coquet catchments. The
379 0.11° CCLM and WRF have better simulation skill than their 0.44° version in all catchments. In contrast, for
380 HIRHAM and RCA, the improvement is only observed in one catchment. For the latter models, there is no
381 added value from increasing the resolution as the simulation processes occurring at higher resolutions than
382 the 0.44° gridbox do not improve the results, possibly due to an inappropriate physical representation. The

383 0.11° CCLM is the best performer in all catchments, except for the Glaslyn where 0.11° HIRHAM has the
384 highest rank.

385 All high-resolution simulations outperform their coarse simulations at the Glaslyn catchment due to
386 the differences between the sizes of the catchment and the different cells. Thus, increasing the RCM
387 resolution increases their simulation skill for catchments with larger elevation variability because the RCMs
388 are able to represent the high-resolution features. In general, increasing the RCM resolution reduces the
389 simulation biases in the upper tail of the distribution, but there are also high-resolution models that
390 consistently overestimate precipitation (e.g. RCA in Figs. 4, S1, S2, S3). The low-resolution models do not
391 simulate the small-sized catchment accurately. In contrast, the flat and large catchments are simulated
392 similarly by both resolutions, showing no added value from increasing RCM resolution.

393 **3.2.1.3. River flow**

394 Now, we evaluate the RCM skill in providing inputs for simulating the river flow in each catchment.
395 In the upper Thames, the 0.11° RCMs overestimate the spring discharge by between 16% and 194% (Fig.
396 6a). Both resolutions underestimate all indices in the Glaslyn catchment (Fig. S4). In the Calder catchment,
397 the 0.44° RCMs underestimate the annual (-9% to -31%) and autumn (-10% to -50%) flows, whereas the
398 0.11° RCMs overestimate the discharge during winter (3% to 63%) and spring (22% to 104%) (Fig. S5a).
399 Also, the Q10 and Q10 annual frequency are underestimated by the 0.44° RCMs (Fig. S5b and c). In the
400 Coquet catchment, the winter mean discharge is underestimated by the 0.44° RCMs by between -7% and
401 -42% and during summer it is overestimated by the 0.11° RCMs by between 2% and 218% (Fig. S6a). In
402 addition, the Q95 is overestimated by the 0.11° simulations.

403 Except for the Glaslyn catchment, the multi-model simulation spread of the flow duration curve
404 (FDC) from both resolutions includes the observed FDC entirely (Fig. 7, first row). For the Glaslyn catchment,
405 both resolutions underestimate the FDC with the 0.11° simulation spread being closer to the observed FDC
406 (Fig. 7d). The 0.44° simulation spread is larger than the 0.11° spread in the Coquet, but smaller in the upper
407 Thames. In the remaining catchments, the spreads of both resolutions are similar.

408 Overall, the maximum monthly NSE values are 0.42 for the Upper Thames (Fig. 6e), 0.22 for the
409 Glaslyn (Fig. S4e), 0.67 for the Calder (Fig. S5e) 0.26 for the Coquet catchment (Fig. S6e), indicating that

410 the best river flow simulation is moderate to poor for all catchments except for the Calder. In contrast, the
411 minimum NSE values are negative in all catchments, implying that there are RCM outputs that generate
412 unreliable river flow simulations even at the monthly times step. Negative NSE values can be a result of
413 river flow overestimation in all indices, for instance 0.11° RCA and HIRHAM in the Calder and Coquet
414 catchments. The Spearman correlation coefficients of the daily river flow are higher for the upper Thames
415 and Calder and smaller for the Glaslyn and Coquet, indicating that the RCMs are able to simulate the daily
416 river flow sequence better on the large and flat sites compared to the small and topographically-complex
417 catchments (Fig. 67f, S4f, S5f and S6f).

418 Comparing their skill in simulating all indices by means of their rank, three 0.11° simulations
419 outperform their 0.44° version in the Upper Thames, five in the Glaslyn, one in the Calder and two in the
420 Coquet catchment (Table 7, last column). Overall, for both resolutions, biases in particular indices are large
421 and the skill of the pairwise indices (NSE, MSE, correlation) is low. The 0.11° simulation biases are
422 consistently smaller than the 0.44° biases only for the Glaslyn catchment due to the difference between the
423 catchment and the 0.44° RCM cell size. However, for this catchment biases are large even for the high-
424 resolution simulations indicating that subgrid processes that result in precipitation increases are not
425 represented by the models. Only CCLM gives better simulation skill for its high-resolution in all catchments.

426 **3.2.2. Bias-corrected RCM simulations**

427 **3.2.2.1. Temperature**

428 Bias-correction reduces the mean and percentile biases by construction (Figs. 3e,f,g,h). Thus, the
429 skill of all RCMs becomes similar in all catchments, as expected. Overall, the larger distribution biases are
430 for the 1st and 99th temperature percentiles, with biases lower than 1°C (Figs. 2, i). Even though these
431 percentiles have the largest biases after bias correction, as may be expected the biases are smaller than
432 those of the uncorrected RCMs. QM does not improve the daily sequence simulation. As a consequence,
433 there is only a slight change in the Pearson correlation coefficient of the daily time series (Figs. 2, iii).

434 3.2.2.2. Precipitation

435 3.2.2.2.1. Gamma distribution QM

436 The skill of both RCM resolutions becomes similar after application of GQM. Nevertheless, biases
437 are not reduced for the 95th percentile, SDII ratio, wet spell length, R95p and R20 in the Upper Thames, for
438 RX1day in the Calder and for the SDII ratio in the Coquet catchment. These indices evaluate the extremes,
439 which are inflated by the correction method (Cannon et al., 2015), and the precipitation intensity.

440 Considering the indices that are not based on the distribution, the Spearman correlation slightly
441 increases after GQM (Figs. 4c, S1c, S2c and S3c) whereas for the MSE the multi-model ensemble bias is
442 reduced, but there are cases when the biases of individual RCMs increase (Figs. 4c, S1c, S2c and S3c).
443 The same happens for the wet and dry spell lengths (Figs. 4b, S1b, S2b and S3b) and RX1day (Figs. 4c,
444 S1c, S2c and S3c). The multi-model bias spread from both resolutions is similar and smaller than 1 mm/day
445 up to the 90th percentile in all catchments (Fig. 5, second row). Above the 90th percentile, the spread of both
446 resolutions increases exponentially. The bias spread in the extremes is larger for the Glaslyn catchment
447 possibly as a consequence of the bias magnitude of the original uncorrected simulation (Fig. 5e).

448 3.2.2.2.2. Double Gamma distribution QM

449 After applying the DGQM method, the skill with respect to distribution-based indices from all RCMs
450 at both resolutions becomes similar. The biases for most distribution-based indices are reduced compared
451 to both uncorrected and GQM. In all catchments, the biases are lower than 1 mm/day below the 99th
452 percentile after which biases increase. Thus, DGQM reduces the percentile biases in all catchments
453 compared to GQM. For the 90th precipitation percentile the DGQM approach increases the biases in all
454 catchments because at this percentile the method segments the precipitation distribution, generating an
455 increment in the bias. Nevertheless, this increase is approximately ± 1 mm/ day in all catchments except
456 the Glaslyn. Additionally, the simulation bias spread of both resolutions is similar for all catchments, as
457 expected (Fig. 5, last row).

458 For the extreme and precipitation intensity measures, DGQM reduces the biases compared to GQM
459 except for the RX1day and SDII ratio in the Upper Thames, R20 in the Glaslyn, R10 in the Calder and the
460 SDII ratio in the Coquet catchment. The simulation skill of the uncorrected models and the GQM and DGQM

461 approaches is similar in all catchments for the Spearman daily correlation coefficient. Overall, the DGQM
462 provides outputs with smaller biases for most of the indices compared to the uncorrected and GQM
463 simulations.

464 **3.2.2.3. River flow**

465 **3.2.2.3.1. Gamma distribution QM**

466 River flow is simulated using the GQM precipitation and temperature as drivers. GQM decreases
467 the bias of all indices in every catchment, except for the Q10 in the upper Thames catchment (Fig. 6c). The
468 bias-corrected FDC simulation spread decreases for both resolutions in all catchments (Fig. 7, second row).
469 The observed FDC is completely included within the spread of both resolutions showing a good simulation
470 of the entire distribution.

471 From the pairwise indices, the skill of the multi-model ensemble improves for the monthly NSE (Fig.
472 6e) and the spread of the daily MSE is reduced in most cases. However, GQM can result in negative NSE
473 values for some models that had positive values when these were not bias-corrected (e.g. 0.44° RACMO
474 and HIRHAM in the Upper Thames). The Spearman correlation of daily time series increases slightly in all
475 cases (Fig. 6f, S4f, S5f and S6i).

476 **3.2.2.3.2. Double Gamma distribution QM**

477 The DGQM approach decreases the biases for all the distribution-based indices compared to both
478 uncorrected and GQM with the exception of Q95 for the Glaslyn catchment. Considering the non-
479 distribution-based indices, the NSE and MSE are not improved for the Coquet catchment. Even though the
480 biases are reduced, the simulation skill among all RCMs does not become similar for specific cases with
481 indices involving the extremes and the pairwise simulation (e.g. the Q10 annual frequency, Q10 and NSE
482 for the Upper Thames, Fig. 6b,c and e). Overall, the daily MSE and monthly NSE simulation skill improves
483 compared to the GQM approach. Thus implying that the river flow simulation skill is better when using the
484 DGQM. By construction of the bias correction method, the FDC simulation spread of both resolutions is
485 similar in shape and amplitude (Fig. 7, bottom row). Compared to GQM, the DGQM simulation spread is
486 further reduced.

487 The Spearman correlation coefficient of the daily river flow time series increases slightly with not a
488 large difference compared to the GQM simulations. Overall, applying the DGQM approach results in smaller
489 biases compared to the GQM, in specific for the simulations of extremes and the monthly sequence.

490 **4. Discussion**

491 Regarding our first research question, as to whether the relative performance of the high- resolution
492 simulations is better than that of the lower-resolution simulations, the results show that the high-resolution
493 RCMs consistently have a better simulation skill for climate and river flow only in the Glaslyn catchment.
494 This is mainly because the size of this catchment is smaller than the 0.44° RCM cell, and it has a complex
495 topography and high precipitation. As a consequence, the skill of the 0.44° simulations in reproducing the
496 local physical features of this catchment is not good. For the other catchments, all of which are larger in
497 size and with less complex topography and less precipitation, both resolutions have a similar performance.
498 Similar results were obtained for the Upper Danube using HIRHAM at resolutions of 50 km x 50 km and 12
499 km x 12 km (Dankers et al., 2007). Only the skill of CCLM improved when using the high-resolution version.
500 Kotlarsky et al. (2014) found that CCLM also gave good results when simulating the mean, seasonal and
501 95th percentile of precipitation over the British Isles. In our study, the remaining RCMs did not improve their
502 simulation skill, implying that the high-resolution versions of these models do not accurately represent
503 processes occurring at higher resolutions.

504 The performance of the two RCM resolutions at simulating temperature was clearly linked to the
505 topographic characteristics of the study catchments. In the upper Thames and Calder catchments, which
506 have relatively flat topography, we found that there is no clear added value from the uncorrected high-
507 resolution RCMs; however, in the topographically-complex Glaslyn and Coquet catchments, all 0.11°
508 simulations outperformed their 0.44° version. These findings are similar to that of Onol et al. (2012) and
509 Tolika et al. (2016) and it is likely that they can be attributed to the difference in elevation from the grid cells
510 of the observations and models, and the lack of representation of the spatial variability. Increases in the
511 simulation skill of local climate when using higher-resolution simulations have been reported before,
512 particularly for mountainous regions (Evans et al., 2013; Larsen et al., 2013; Tolika et al., 2016).

513 The uncorrected 0.11° simulations largely underestimate the precipitation and river flow
514 observations of the Glaslyn catchment, mainly due to the catchment's topographic complexity and high
515 levels of precipitation. Similar results for the Euro-CORDEX RCMs have been obtained for precipitation in
516 other regions with complex topography (e.g. Casanueva et al., 2016; Prein et al., 2015; Torma et al., 2015).
517 For the remaining catchments, the multi-model simulation spread of the simulations of both resolutions
518 includes the observed FDC, indicating that the models are able to provide useful simulations that resemble
519 the observed river flow. However, the simulation spread can be large; deviations in the annual mean river
520 flow reach almost 100% for some RCMs. Individual uncorrected RCMs have small biases and satisfactory
521 simulations of the river flow (e.g. 0.11° CCLM in the Calder and Coquet catchments), but there are also
522 RCMs that are not able to provide useful simulations. For example, the 0.11° RCA had the largest
523 precipitation and river flow biases in most indices for all catchments. In contrast, all the bias-corrected RCM
524 simulations are closer to the observed climate and river flow. Furthermore, the simulation skill from all bias-
525 corrected RCMs at both resolutions becomes similar and as a result, the simulation spread of the multi-
526 model ensemble is reduced compared to the uncorrected simulations, providing a smaller range of possible
527 scenarios.

528 Our results show that uncorrected RCMs provide river flow simulations that have too much spread
529 to be able to be used for impact studies (also stated by Kay et al., 2015; Cloke et al., 2013). Both resolutions
530 have a similar performance when simulating the seasonal mean river flow as there are biases from both
531 resolutions. However, certain high-resolution models tend to overestimate the seasonal flow largely for
532 most of the catchments and seasons (e.g. RCA in all catchments and HIRHAM in the Coquet and Calder
533 catchments). In contrast, the low-resolution CCLM underestimates river flow for all seasons and catchments.
534 For the medium-sized Calder catchment, individual models have different biases per season but the multi-
535 model ensemble mean shows a consistent underestimation for high-resolution models and underestimation
536 of river flow for the low-resolution modes. This is not distinguished in the larger Upper Thames nor in the
537 Coquet catchment. Similar to the annual mean flow, both resolutions underestimate the seasonal flow in
538 the Glaslyn catchment. In comparison, all the bias-corrected RCMs simulate the river flow much closer to
539 the observed flows and reduce the simulation spread, thus providing plausible inputs for impact studies.

540 Finally, to answer our last research question, we evaluate the simulation skill of DGQM compared
541 to GQM. Using four catchments with different characteristics, the DGQM provides a better simulation of the
542 river flow characteristics compared to the QGM approach, with a higher improvement for the simulation of
543 extremes and the monthly sequence. The GQM systematically reduces the precipitation bias up to the 90th
544 percentile, but exponentially increases the bias above this percentile. Therefore, to capture the properties
545 of extremes, we suggest using the DGQM with the 90th percentile as segmentation threshold in contrast to
546 Yang et al. (2010) who divided the distribution at the 95th percentile. Based on our results, the DGQM
547 reduces the precipitation and river flow biases of most indices compared to the commonly used GQM. This
548 is particularly relevant for the analysis of extreme precipitation and high flows as the GQM is usually
549 employed in flood analysis (e.g. Cloke et al., 2013) and river flow projections (e.g. Prudhomme et al., 2013).
550 In addition, the DGQM reduces the ensemble spread more than the GQM, without introducing much extra
551 complexity. However, no bias correction method will remove all biases. Thus, the selection of the method
552 depends on the requirements of each study (Nguyen et al., 2017) and it should be tested to evaluate
553 whether the benefits justify their calculation complexities.

554 Ideally, RCMs should not require post-processing techniques to provide simulations which can be
555 used with confidence (Ehret et al., 2012). However, our results demonstrate large biases for various
556 diagnostic indices for the reanalysis-driven RCMs. Particular RCMs provide plausible river flow simulations,
557 for instance, 0.11° CCLM for the Calder catchment when assessing the annual and seasonal means, low
558 flows, high flow occurrence and pairwise indices. However, the RCM simulation skill is catchment-
559 dependent. Thus, at the moment, bias correction seems to be the best approach to reduce the ensemble
560 spread and its biases. Nevertheless, bias correction methods should be used carefully for the analysis of
561 future projections (Cloke et al., 2013) as bias correction cannot correct fundamental problems from the
562 original climate model (Maraun and Widmann, 2015; Maraun et al., 2017) and the spread of the bias-
563 corrected simulations might not reflect the total real uncertainty. Climate research is focusing on
564 determining the causes behind the biases (e.g. Addor et al., 2016) and improving the simulation of the
565 processes (e.g. Zittis et al., 2017; Meredith et al., 2015). For instance, convection permitting models seek
566 to improve the simulation of precipitation extremes (Tölle et al., 2017; Gutjahr et al., 2016). However, the

567 computational cost of developing such models is large and, as a consequence, the simulation length is
568 short and the availability of GCM-RCM projections is low.

569 By analysing four catchments with different characteristics, we evaluate the RCM simulation skill in
570 different contexts. Our results suggest that the small size and the high precipitation (e.g. Glaslyn catchment)
571 are the main factors related to the better simulation skill from the high-resolution RCMs over the low-
572 resolution models for the simulation of river flow. The importance of topographical complexity and other
573 characteristics for the simulation outputs is secondary. This is highlighted by the results of the medium-
574 sized Coquet catchment, for which both resolutions have similar simulation skill even with its complex
575 topography. Although the hydrological model used (HEC-HMS) was chosen as it has been used before in
576 assessment of climate change impacts (e.g. Babel et al., 2014; Azmat et al., 2015) and acknowledging that
577 there are a diversity of methods used to simulate the hydrological processes, we note that our results are
578 unlikely to substantially change when using other hydrological model. We support our statement as we
579 assess the performance of the different resolutions by evaluating the RCM outputs as well as the
580 hydrological model outputs, both giving similar results. An assessment of the hydrological model uncertainty
581 is beyond the scope of this study, but will be the subject of future work.

582 **5. Conclusions**

583 This study provides information on the added value from increasing RCM resolution and bias
584 correction techniques for the simulation for river flow. Previous studies have assessed the improvement in
585 the simulation skill of climate variables due to an increase in the RCM resolution, but this might not
586 guarantee an improvement in the simulation of the river flow parameters that are relevant for impact studies.
587 We conducted a comprehensive analysis on how the uncorrected and bias-corrected RCM outputs drive
588 the simulations of river flow at high and low resolutions. Each RCM used here has the same
589 parameterisation, domain and driving data at both resolutions, and therefore the comparison only evaluates
590 the effect of increasing its resolution. We analysed four catchments located at different latitudes within
591 Great Britain. These catchments vary in climate (e.g. precipitation ranging from 2900 mm yr⁻¹ to 762 mm
592 yr⁻¹), physical characteristics (flat and complex topographies, areas ranging from 69 km² to 1616 km²), land
593 use (varying from urban-dominant to agricultural and natural areas) and hydrological characteristics (e.g.

594 annual mean river flow ranging from $15.3 \text{ m}^3 \text{ s}^{-1}$ to $5.8 \text{ m}^3 \text{ s}^{-1}$). We applied a detailed assessment of the
595 simulation skill of the climate and hydrological models using a set of indices relevant for the analysis of
596 different impacts.

597 We show that the uncorrected 0.11° RCMs only showed better skill in simulating precipitation and
598 river flow in the small catchment. This is because the spatial resolution of the 0.44° models is four-times
599 larger than the catchment size, whereas one cell of the 0.11° model is similar in area to the catchment.
600 Nevertheless, the high-resolution simulations are not able to accurately represent the complex topography
601 of this catchment and do not resolve local processes, underestimating the observed precipitation and the
602 entire FDC. In Australia, Parana Manage et al. (2016) also found that the averaging of topography of gridded
603 outputs influences on the accurate simulation of rainfall.

604 Both resolutions capture the temperature and precipitation distribution, as well as the FDC, for the
605 remaining sites. Thus, in principle, the simulations could be used for climate change assessments.
606 Nevertheless, for most of the indices, the multi-model variability is large (e.g. the mpe of the annual mean
607 river flow simulation ranges from 198% to -31% in the Upper Thames, with an average of 49%), making
608 any interpretation difficult in practice. Only one RCM (CCLM) improves the river flow simulation when using
609 its high-resolution version in all catchments, implying that the remaining models do not simulate the relevant
610 high-resolution processes accurately as there is no consistent difference between their high and low
611 resolution versions. Therefore, there is no added value from using the high-resolution RCMs in those
612 catchments for the assessment of river flow impacts.

613 Bias-correction reduces the distribution-based biases for all RCMs and resolutions by construction.
614 Thus, the bias-corrected high- and low-resolution RCMs have similar simulation skill for the distribution-
615 based indices. There is also less spread from the ensemble simulation of precipitation and river flows (e.g.
616 the mpe of the annual mean river flow simulations for the Upper Thames ranges from -1% to 16% when
617 corrected using DGQM, with an average of 7%). Nevertheless, daily pairwise indices, which assess the skill
618 of the model when simulating the observed time series, are not improved by bias correction. However, the
619 monthly NSE results indicate that bias correction can improve the pairwise simulation on monthly
620 timescales. Overall, correcting the RCMs to the local temperature and precipitation provides a reduction of

621 the ensemble spread, making the outputs more useful for the analysis of impacts. Nevertheless, it should
622 be considered that the ensemble spread of uncorrected and corrected models can underestimate the true
623 simulation uncertainty.

624 In comparison to GQM, DGQM provides a larger reduction in the simulation biases for precipitation
625 and river flow. The main difference between both methods is the greater correction from DGQM for
626 precipitation extremes (95th percentile, R10, R20, R95p) and high flows (Q10 and Q10 annual frequency).
627 The monthly NSE consistently shows an improvement in the simulation skill of RCMs that are corrected
628 using DGQM. Overall, for most of the RCMs and considering the results from all indices, the DGQM
629 outperforms GQM.

630 Our study shows that an increase in RCM resolution does not always imply a better simulation of
631 hydrological impacts, especially for large catchments. In contrast, small catchments with complex
632 topography are still difficult to be simulated accurately by high-resolution models, concurring with Dankers
633 et al (2007). The uncorrected RCM ensemble generally provides a large spread which makes it difficult to
634 use for impact assessment. Similar outcomes have been obtained for other regions, for example Australia
635 (Lockart et al. 2016), Canada and China (Wang et al. 2019). Bias-correction provides an alternative to
636 reduce the biases and multi-model spread, making decision-making easier. From the methods evaluated
637 here, DGQM reduces most of the RCM biases without much more complexity added to the bias-correction
638 method employed when using GQM. However, and agreeing with Cloke et al. (2013), Huang et al. (2014)
639 and Lockart et al. (2016), the bias-corrected outputs should be used carefully when evaluating changes in
640 very extreme flows as the correction inflates the simulated extremes. Compared to previous studies, we
641 can state that our results are robust as we included a larger number of RCMs with different
642 parameterisations for our analysis.

643 Whilst effective, bias correction adds extra uncertainty to the analysis chain (Cloke et al., 2013;
644 Rummukainen, 2016). Therefore, it must be used with consideration of its limitations: dependence on the
645 training period (Lafon et al., 2013), assumption of temporal stability of the correction function (Chen et al.,
646 2015), issues of sub-grid variability and inflation of variance (Maraun, 2013), inter-variable consistency
647 (Wilcke et al., 2013), spatial representation over complex terrain (Maraun and Widmann, 2015) and biases

648 from the driving data (Maraun et al., 2017). The extent to which the climate change signal is altered must
 649 also be considered (Maraun, 2013; Velázquez et al., 2015) along with the possibility that bias correction
 650 can produce larger biases for extremes than for the mean (Huang et al., 2014). Additionally, we
 651 acknowledge that using different data to drive the RCMs used in this study, for instance a GCM, could give
 652 different results, as could the use of a different hydrological model.

653 Our results can provide an insight on whether RCMs of high(er) resolution improve the simulation skill.
 654 These can be useful for regions of similar characteristics where high(er)-resolution RCMs have not been
 655 developed yet and would require considerable time and effort to be produced. If used, bias-correction
 656 methods should be tested for the specific analysis that will be performed. This study provided different
 657 methods to perform this testing for the different RCMs and bias-correction methods for climatology and
 658 hydrology.

659

660 Acknowledgements

661 Pastén-Zapata received funding from Consejo Nacional de Ciencia y Tecnología (Conacyt) and Secretaría
 662 de Educación Pública (SEP) during the development of this study. We also thank the Euro-CORDEX
 663 initiative for providing the climate simulations and the US Army Corp of Engineers for making the HEC-
 664 HMS model available. We are also grateful for the historical observational datasets provided by the UK's
 665 CEH and NERC. Finally, we thank two anonymous reviewers for their comments regarding this paper.

666

667

668

669 Tables

670

Table 1. Characteristics of the study sites

	Upper Thames	Glaslyn	Calder	Coquet
Area (km ²)	1616	69	316	346
Maximum elevation (masl ¹)	330	1080	556	775
Minimum elevation (masl ¹)	52	30	40	71
Mean annual precipitation (mm/year)	762	2957	1251	968
Mean annual temperature (°C)	9.7	8.1	8.4	7.4
Mean annual PET (mm/yr)	522	477	486	473
Mean annual river flow (m ³ /s)	15.3	5.8	8.8	6.1

Precipitation 90th percentile (mm/day)	6.7	24.4	10.3	7.7
Precipitation 95th percentile (mm/day)	10.2	34.2	14.8	11.9
² Q10 (m ³ /s)	34.8	13.5	19.9	12.4
³ Q95 (m ³ /s)	1.90	0.55	1.99	0.84

671 ¹ Meters above the sea level

672 ² River flow that is exceeded for 10% of the daily river flow time series

673 ³ River flow that is exceeded for 95% of the daily river flow time series

674

675

676

Table 2. RCMs used in this study

RCM	Institute	Period	Reference
CCLM-CLMCOM	Brandenburg University of Technology (BTU)	1989-2008	Böhm et al., 2006; Rockel et al., 2008
HIRHAM 5	Danish Meteorological Institute (DMI)	1989-2008	Christensen et al., 1998
RACMO22E	Royal Netherlands Meteorological Institute (KNMI)	1979-2008	Van Meijgaard et al., 2012
RCA4	Swedish Meteorological and Hydrological Institute (SMHI)	1984-2008	Samuelsson et al., 2011
WRF 3.3.1	Institute Pierre Simon Laplace (IPSL) and Institute National de l'Environnement Industriel et des Risques (INERIS)	1989-2008	Skamarock et al., 2008

677

Index	Description	Performance measure
Precipitation		
95 th percentile	A measure of very extreme events: 95 th percentile of daily precipitation	Bias (mm/day)
90 th percentile	A measure of extreme events: 90 th percentile of daily precipitation	Bias (mm/day)
50 th percentile	50 th percentile of daily precipitation	Bias (mm/day)
25 th percentile	25 th percentile of daily precipitation	Bias (mm/day)
^a Wet spell length	Mean wet spell length for a given month of the year	Bias (days)
^a Dry spell length	Mean dry spell length for a given month of the year	Bias (days)
^a Annual mean precipitation	Annual accumulated precipitation	Mean percentage error
^a Monthly mean precipitation	Accumulated precipitation for a given month of the year	Mean percentage error
^b Relative daily MSE	Mean daily square error, shown as ratio to the largest MSE result (considering both corrected and uncorrected RCMS)	MSE (ratio)
^b Spearman correlation coefficient	Spearman correlation coefficients between the daily simulated and observed time series	Index
^a Maximum one day precipitation (RX1day)	Maximum one-day precipitation for a given month of the year	Mean percentage error
^a Simple Daily Intensity Index (SDII)	Ratio of the annual total precipitation to the number of wet days (≥ 1 mm) in all years	Index
^a Number of heavy precipitation days (R10)	Mean number of days with precipitation ≥ 10 mm within a year	Bias (days)
^a Number of very heavy precipitation days (R20)	Mean number of days with precipitation ≥ 20 mm within a year	Bias (days)
^a Very wet days (R95p)	Mean annual accumulated precipitation from days > 95 th percentile in all years	Mean percentage error
Temperature		
^a Annual mean temperature	Annual mean temperature over the validation period	Mean percentage error
^a Monthly mean temperature	Monthly mean temperature	Mean percentage error
99 th percentile of daily mean temperature	99 th percentile of the daily mean temperature	Bias ($^{\circ}$ C/day)
1 st percentile of daily mean temperature	1 st percentile of the daily mean temperature	Bias ($^{\circ}$ C/day)
^b Pearson correlation coefficient	Pearson correlation coefficient between the daily RCM and observation time series	Index
River Flow		
Q10	A measure of high flows: river flow that is exceeded for 10% of the daily river flow time series	Bias (m ³ /s)
Q95	A measure of low flows: river flow that is exceeded for 95% of the daily river flow time series	Bias (m ³ /s)
^a Annual Q10 frequency	Mean number of days for which the observed Q10 is exceeded within a year	Bias (days)
^a Annual mean river flow	Annual mean daily river flow over the validation period	Mean percentage error
^a Winter (DJF) mean river flow	Winter mean daily river flow over the validation period	Mean percentage error
^a Spring (MAM) mean river flow	Spring mean daily river flow over the validation period	Mean percentage error
^a Summer (JJA) mean river flow	Summer mean daily river flow over the validation period	Mean percentage error
^a Autumn (SON) mean river flow	Autumn mean daily river flow over the validation period	Mean percentage error
^b Monthly NSE	Monthly Nash Sutcliffe Efficiency index	Index
^b Relative daily MSE	Mean daily square error, shown as ratio to the largest MSE result (considering both corrected and uncorrected RCMS)	MSE (ratio)
^b Spearman correlation coefficient	Spearman correlation coefficient between the daily simulated and observed time series	Index

^a Estimated using the long term mean (one value over the entire series)

^b Estimated considering the time series values (one value per time step)

680
681

Table 4. Indices from the calibration and validation of the hydrological models

Catchment	Step	Period	Daily NSE	Q10 bias		Q95 bias	
				(m ³ /s)	(%)	(m ³ /s)	(%)
Upper Thames	Calibration	1986-2010	0.70	-2.1	-6	-0.45	-25
	Validation	1961-1985	0.57	1.5	5	-0.90	-44
Glaslyn	Calibration	1991-2010	0.78	1.0	8	-0.07	-11
	Validation	1971-1990	0.78	0.7	5	-0.03	-6
Calder	Calibration	1994-2010	0.62	1.5	8	-0.31	-16
	Validation	1976-1993	0.60	1.3	7	-0.24	-12
Coquet	Calibration	1992-2010	0.63	1.3	11	-0.24	-27
	Validation	1973-1991	0.52	-0.6	-5	-0.25	-31

682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707

708
709

Table 5. RCM rank for the temperature indices for each catchment: 1 = best, 10 = worst. The asterisks (*) indicate the resolution with the best simulation skill of each RCM in each catchment

		99th percentile	1st percentile	Annual mean	Monthly mean	Correlation	Average score	Ranking	
Upper Thames	0.11°CCLM	10	7	2	9	1	5.8	6	*
	0.11°HIRHAM	3	9	3	5	6	5.2	5	
	0.11°RACMO	2	8	9	7	4	6.0	7	
	0.11°RCA	7	5	10	10	5	7.4	10	
	0.11°WRF	4	1	5	4	8	4.4	2	*
	0.44°CCLM	9	10	1	8	2	6.0	7	
	0.44°HIRHAM	1	6	4	3	9	4.6	3	*
	0.44°RACMO	5	4	7	2	3	4.2	1	*
	0.44°RCA	8	2	6	1	7	4.8	4	*
0.44°WRF	6	3	8	6	10	6.6	9		
Glaslyn	0.11°CCLM	9	2	4	3	1	3.8	3	*
	0.11°HIRHAM	7	6	2	4	7	5.2	5	*
	0.11°RACMO	3	7	1	1	4	3.2	1	*
	0.11°RCA	2	4	3	2	6	3.4	2	*
	0.11°WRF	4	8	5	6	10	6.6	7	*
	0.44°CCLM	10	1	6	5	2	4.8	4	
	0.44°HIRHAM	8	3	8	7	9	7.0	8	
	0.44°RACMO	5	5	7	8	3	5.6	6	
	0.44°RCA	6	9	9	9	5	7.6	9	
0.44°WRF	1	10	10	10	8	7.8	10		
Calder	0.11°CCLM	9	7	8	8	1	6.6	7	
	0.11°HIRHAM	5	9	7	7	5	6.6	7	
	0.11°RACMO	8	10	10	10	4	8.4	9	
	0.11°RCA	10	8	9	9	6	8.4	9	
	0.11°WRF	7	3	1	4	8	4.6	4	*
	0.44°CCLM	6	6	6	5	2	5	5	*
	0.44°HIRHAM	4	2	2	1	9	3.6	2	*
	0.44°RACMO	2	4	5	2	3	3.2	1	*
	0.44°RCA	3	1	4	3	7	3.6	2	*
0.44°WRF	1	5	3	6	10	5	5		
Coquet	0.11°CCLM	9	2	2	3	2	3.6	3	*
	0.11°HIRHAM	1	3	3	2	5	2.8	1	*
	0.11°RACMO	3	7	9	7	4	6.0	5	*
	0.11°RCA	7	6	8	4	6	6.2	6	*
	0.11°WRF	5	1	1	1	8	3.2	2	*
	0.44°CCLM	4	4	7	5	1	4.2	4	
	0.44°HIRHAM	10	8	5	6	9	7.6	9	
	0.44°RACMO	6	9	6	9	3	6.6	8	
	0.44°RCA	2	5	10	8	7	6.4	7	
0.44°WRF	8	10	4	10	10	8.4	10		

710

711 Table 6. RCM rank for precipitation: 1-best, 10-worst. The asterisks (*) indicate the resolution with the best simulation skill of each RCM in each catchment

		Pr 95th	Pr 90th	Pr 50th	Pr 25th	Annual Mean	Monthly MSE	Dry Spell	Wet Spell	Monthly Mean	Correl.	SDII	R10	R20	R95p	RX1day	Average score	Ranking	
Upper Thames	0.11°CCLM	8	5	1	2	5	2	1	4	5	1	6	8	4	8	2	4.1	1	*
	0.11°HIRHAM	7	4	4	3	3	5	6	6	1	3	5	7	6	3	7	4.7	4	
	0.11°RACMO	3	2	9	8	7	3	4	5	4	10	9	3	5	5	9	5.7	8	
	0.11°RCA	10	10	10	10	10	10	10	10	10	6	2	10	10	10	8	9.1	10	
	0.11°WRF	1	1	6	7	6	8	7	3	8	5	7	1	3	1	3	4.5	2	*
	0.44°CCLM	9	9	2	1	8	4	3	8	7	2	4	9	8	9	1	5.6	7	*
	0.44°HIRHAM	5	6	3	5	2	7	5	9	3	7	3	4	1	4	5	4.6	3	*
	0.44°RACMO	4	3	5	6	4	1	2	2	2	9	8	5	7	6	6	4.7	4	*
	0.44°RCA	2	8	8	4	9	9	9	1	9	4	1	2	1	2	4	4.9	6	*
0.44°WRF	6	7	7	9	1	6	8	7	6	8	10	6	9	7	10	7.1	9		
Glaslyn	0.11°CCLM	5	5	8	2	5	5	6	5	5	1	5	5	5	5	5	4.8	5	*
	0.11°HIRHAM	1	1	6	5	1	3	5	3	2	3	1	3	1	1	1	2.5	1	*
	0.11°RACMO	3	3	3	9	3	1	3	8	3	2	3	2	3	2	4	3.5	3	*
	0.11°RCA	2	2	2	10	2	2	8	6	1	6	2	1	2	3	2	3.4	2	*
	0.11°WRF	4	4	1	6	4	4	4	4	4	7	4	4	4	4	3	4.1	4	*
	0.44°CCLM	10	9	10	3	9	9	9	9	9	5	9	9	9	8	7	8.3	9	
	0.44°HIRHAM	9	10	9	1	10	10	10	10	10	9	7	10	10	9	9	8.9	10	
	0.44°RACMO	7	7	4	7	7	7	2	1	7	4	10	7	7	7	8	6.1	7	
	0.44°RCA	8	8	7	4	8	8	7	7	8	8	8	8	8	10	10	7.8	8	
0.44°WRF	6	6	5	8	6	6	1	2	6	10	6	6	6	6	6	5.7	6		
Calder	0.11°CCLM	1	2	2	1	1	1	2	3	1	1	1	2	2	2	9	2.1	1	*
	0.11°HIRHAM	10	10	8	5	9	9	7	8	9	5	7	9	10	10	10	8.4	9	
	0.11°RACMO	2	1	9	9	3	5	5	9	4	4	4	1	1	1	3	4.1	2	*
	0.11°RCA	9	9	10	10	10	10	10	10	10	6	3	10	9	9	1	8.4	9	
	0.11°WRF	3	3	6	4	6	8	6	5	7	8	2	4	3	3	8	5.1	5	*
	0.44°CCLM	6	7	4	2	8	3	4	4	8	2	5	7	4	6	2	4.8	3	
	0.44°HIRHAM	4	4	1	3	7	4	9	6	6	7	6	3	5	4	5	4.9	4	*
	0.44°RACMO	8	8	7	7	5	2	3	1	3	3	10	8	8	8	6	5.8	7	
	0.44°RCA	7	6	3	6	4	7	8	2	5	9	8	6	7	7	7	6.1	8	*
0.44°WRF	5	5	5	8	2	6	1	7	2	10	9	5	6	5	4	5.3	6		
Coquet	0.11°CCLM	4	5	1	1	2	1	1	3	1	1	3	4	2	1	2	2.1	1	*
	0.11°HIRHAM	6	9	9	7	9	9	9	7	9	5	1	7	5	6	4	6.8	8	
	0.11°RACMO	5	3	6	8	1	3	7	5	2	4	9	5	4	5	5	4.8	4	*
	0.11°RCA	10	10	10	10	10	10	10	10	10	9	7	10	9	10	1	9.1	10	
	0.11°WRF	2	1	5	3	3	6	5	6	3	7	5	2	1	2	3	3.6	2	*
	0.44°CCLM	7	6	4	2	8	4	3	8	7	2	3	6	7	7	6	5.3	6	
	0.44°HIRHAM	3	2	8	9	4	5	8	2	4	8	1	1	3	4	8	4.7	3	*
	0.44°RACMO	8	7	3	4	6	2	2	4	6	3	9	8	10	9	10	6.1	7	
0.44°RCA	1	4	7	5	5	8	6	1	5	6	7	3	6	3	7	4.9	5	*	
0.44°WRF	9	8	2	6	7	7	4	9	8	10	5	9	8	8	9	7.3	9		

712 Table 7. RCM rank for river flow: 1-best, 10-worst. The asterisks (*) indicate the resolution with the best simulation skill of each RCM in each catchment

		Annual mean	Winter mean	Spring mean	Summer mean	Autumn mean	Monthly NSE	Daily MSE	Correl.	Q10	Q10 frequency	Q95	Average score	Rank	
Upper Thames	0.11° CCLM	1	2	2	3	3	1	1	4	1	1	1	1.8	1	*
	0.11° HIRHAM	3	4	4	2	1	3	3	1	3	3	4	2.8	2	*
	0.11° RACMO	8	9	8	6	9	9	8	7	8	8	9	8.1	9	
	0.11° RCA	10	10	10	10	10	10	10	10	10	10	10	10.0	10	
	0.11° WRF	7	1	6	9	5	6	6	8	5	6	8	6.1	6	*
	0.44° CCLM	4	7	1	4	7	2	2	3	6	4	3	3.9	4	
	0.44° HIRHAM	2	5	3	1	2	4	5	5	2	2	2	3.0	3	
	0.44° RACMO	5	6	5	5	6	5	4	2	4	5	7	4.9	5	*
	0.44° RCA	9	8	9	7	8	8	9	6	9	9	6	8.0	8	*
0.44° WRF	6	3	7	8	4	7	7	9	7	7	5	6.4	7		
Glaslyn	0.11° CCLM	5	5	5	6	6	5	5	4	5	5	8	5.36	5	*
	0.11° HIRHAM	1	1	1	4	1	1	3	2	1	1	4	1.82	1	*
	0.11° RACMO	2	2	2	3	3	2	1	1	2	2	2	2	2	*
	0.11° RCA	3	3	3	1	2	3	2	6	3	3	1	2.73	3	*
	0.11° WRF	4	4	4	2	4	4	4	5	4	4	3	3.82	4	*
	0.44° CCLM	9	9	10	10	10	9	9	8	10	8	9	9.18	9	
	0.44° HIRHAM	10	10	9	9	9	10	10	9	9	8	10	9.36	10	
	0.44° RACMO	7	6	7	8	7	7	7	3	7	7	6	6.55	7	
	0.44° RCA	8	8	8	7	8	8	8	10	8	10	7	8.18	8	
0.44° WRF	6	7	6	5	5	6	6	7	6	6	5	5.91	6		
Calder	0.11° CCLM	2	2	6	4	5	2	1	2	1	1	1	2.45	2	*
	0.11° HIRHAM	9	10	9	8	9	9	9	4	9	9	9	8.55	9	
	0.11° RACMO	6	6	7	6	6	5	6	3	6	8	8	6.09	6	
	0.11° RCA	10	9	10	10	10	10	10	9	10	10	10	9.82	10	
	0.11° WRF	7	8	8	9	2	8	8	6	7	7	7	7	8	
	0.44° CCLM	8	5	5	7	8	6	3	5	8	6	6	6.09	6	
	0.44° HIRHAM	5	4	3	5	7	4	5	10	5	5	4	5.18	5	*
	0.44° RACMO	3	3	2	1	1	1	2	1	3	2	2	1.91	1	*
	0.44° RCA	4	7	4	3	3	7	7	8	4	4	5	5.09	4	*
0.44° WRF	1	1	1	2	4	3	4	7	2	3	3	2.82	3	*	
Coquet	0.11° CCLM	1	5	1	1	6	4	5	1	4	1	1	2.73	1	*
	0.11° HIRHAM	9	6	9	9	8	9	9	5	8	9	9	8.18	9	
	0.11° RACMO	7	1	7	7	5	6	3	6	2	2	8	4.91	5	
	0.11° RCA	10	10	10	10	10	10	10	10	10	10	10	10	10	
	0.11° WRF	2	2	2	5	2	2	6	4	1	3	5	3.09	2	*
	0.44° CCLM	8	9	8	4	9	7	2	2	9	8	2	6.18	7	
	0.44° HIRHAM	6	3	4	8	7	8	8	9	5	7	7	6.55	8	*
	0.44° RACMO	3	7	3	2	3	1	1	3	6	6	4	3.55	3	*
	0.44° RCA	5	4	6	6	4	5	7	7	3	4	6	5.18	6	*
0.44° WRF	4	8	5	3	1	3	4	8	7	5	3	4.64	4		

714 **References**

- 715 Addor, N., Rohrer, M., Furrer, R., Seibert, J., 2016. Propagation of biases in climate models from synoptic to the
716 regional scale: Implications for bias adjustment. *J. Geophys. Res. Atmos.*, 121.
717 doi:10.1002/2015JD024040.
- 718 Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop
719 water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9), D05109.
- 720 Arnell, N.W. 2011. Uncertainty in the relationship between climate forcing and hydrological response in UK
721 catchments. *Hydrology and Earth System Sciences*, 15(3), 897-912
- 722 Azmat, M., Laio, F., Poggi, D., 2015. Estimation of water resources availability and mini-hydro productivity in
723 high-altitude scarcely-gauged watershed. *Water Resour Manage*, 29:5037-5054. doi: 10.1007/s11269-
724 015-1102-z.
- 725 Babel, M.S., Bhusal, S.P., Wahid, S.M., Agarwal, A., 2014. Climate change and water resources in the Bagmati
726 River Basin, Nepal. *Theor. Appl. Climatol.* 115,639-654. doi: 10.1007/s00704-013-0910-4.
- 727 Benestad, R.E., 2010. Downscaling precipitation extremes. *Theoretical and Applied Climatology*, 100(1-2), 1-21.
- 728 Böhm, U., Kücken, M., Ahrens, W., Block, A., Hauffe, D., Keuler, K., Rockel, B., Will, A., 2006. CLM-the climate
729 version of LM: Brief description and long-term applications. Tech. Rep., COSMO Newsletter.
- 730 Bucchignani, E., Montesarchio, M., Zollo, A.L. and Mercogliano, P., 2016. High-resolution climate simulations
731 with COSMO-CLM over Italy: performance evaluation and climate projections for the 21st
732 century. *International Journal of Climatology*, 36(2), pp.735-756.
- 733 Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: How
734 well do methods preserve changes in quantiles and extremes? *Journal of Climate*, 28(17), 6938-6959.
- 735 Casanueva, A., Kotlarski, S., Herrera, S., Fernández, J., Gutiérrez, J. M., Boberg, F., Colette, A., Christensen,
736 O. B., Goergen, K., Jacob, D., Keuler, K., Nikulin, G., Teichmann, C., Vautard, R., 2016. Daily
737 precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected
738 high-resolution simulations. *Clim. Dyn.* 1-19. doi: 10.1007/s00382-015-2865-x

- 739 Chen, J., Brissette, F.P. and Lucas-Picher, P., 2015. Assessing the limits of bias-correcting climate model
740 outputs for climate change impact studies. *Journal of Geophysical Research: Atmospheres*, 120(3),
741 1123-1136.
- 742 Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of
743 regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.* 35(20),
744 L20709. doi: 10.1029/2008GL035694.
- 745 Christensen, O.B., Christensen, J.H., Machehauer, B., Botzet, M., 1998. Very high-resolution regional climate
746 simulations over Scandinavia-present climate. *J. Clim.* 11(12), 3204-3229.
- 747 Christerson, B.V., Vidal, J.P., Wade, S.D. 2012. Using UKCP09 probabilistic climate information for UK water
748 resource planning. *J. Hydrol.* 424-425: 48-67. doi: 10.1016/j.hydrol.2011.12.020.
- 749 Cloke, H.L., Wetterhall, F., He, Y., Freer, J.E., Pappenberger, F., 2013. Modelling climate impact on floods with
750 ensemble climate projections. *Q. J. R. Meteorol. Soc.* 139, 282-297. doi: 10.1002/qj.1998.
- 751 Dankers, r., Christensen, O.B., Feyen, L., Kalas, M., de Roo, A., 2007. Evaluation of very high-resolution climate
752 model data for simulating flood hazards in the Upper Danube Basin. *J. Hydrol.*, 347: 319-331. doi:
753 10.1016/j.hydrol.2007.09.055.
- 754 Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A.,
755 Balsamo, G., Bauer, P., Bechtold, P., 2011. The ERA-Interim reanalysis: Configuration and performance
756 of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656), 553-
757 597.
- 758 Di Luca, A., de Elía, R., Laprise, R., 2015. Challenges in the quest for added value of regional climate dynamical
759 downscaling. *Curr. Clim. Change Rep.* 1, 10-21. doi: 10.1007/s40641-015-0003-9.
- 760 Eden, J.M., Widmann, M., Maraun, D., Vrac, M., 2014. Comparison of GCM- and RCM-simulated precipitation
761 following stochastic postprocessing. *J. Geophys. Res. Atmos.* 119, 11040-11053.
762 doi:10.1002/2014JD021732.
- 763 Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. Should we apply bias correction to global
764 and regional climate model data? *Hydrol. Earth Syst. Sci.* 16, 3391-3404. doi: 10.5194/hess-16-3391-
765 2012.

- 766 Evans, J.P., McCabe, M.F., 2013. Effect of model resolution on a regional climate model simulation over
767 southeast Australia. *Clim. Res.* 56:131-145. doi: 10.1007/s00382-012-1415-z.
- 768 Fantini, A., Raffaele, F., Torma, C., Bacer, S., Coppola, E., Giorgi, F., Ahrens, B., Dubois, C., Sanchez, E. and
769 Verdecchia, M., 2016. Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-
770 CORDEX and EURO-CORDEX experiments against high resolution observations. *Climate Dynamics*, 1-
771 24.
- 772 Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., Zahn, M., 2011. Regional climate models add value to
773 global model data. A review and selected examples. *Bull. Amer. Meteor. Soc.* September. doi:
774 10.1175/2011BAMS3061.1.
- 775 Fowler, H.J. and Kilsby, C.G., 2007. Using regional climate model data to simulate historical and future river flows
776 in northwest England. *Clim. Change.* 80: 337-367.
- 777 Gutjahr, O., Schefczyk, L., Reiter, P. and Heinemann, G., 2016. Impact of the horizontal resolution on the
778 simulation of extremes in COSMO-CLM. *Meteorol. Z.*
- 779 Hough, M.N. and Jones, R.J.A., 1997. The United Kingdom Meteorological Office rainfall and evaporation
780 calculation system: MORECS version 2.0-an overview. *Hydrology and Earth System Sciences*
781 *Discussions*, 1(2), pp.227-239.
- 782 Huang, S., Krysanove, V., Hatterman, F. F., 2014. Does bias correction increase reliability of flood projections
783 under climate change? A case study of large rivers in Germany. *Int. J. Climatol.* 34, 3780-3800. doi:
784 10.1002/joc.3945.
- 785 Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M.,
786 Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N.,
787 Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsman, A., Martin, E., van Meijgaard, E.,
788 Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M.,
789 Samuelsson, P., Somot, S., Soussana, J.F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou,
790 P., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact
791 research. *Reg. Environ. Change.* 14,563-578. doi: 10.1007/s10113-013-0499-2.

- 792 Kay, A.L. and Davies, H.N., 2008. Calculating potential evaporation from climate model data: A source of
793 uncertainty for hydrological climate change impacts. *Journal of Hydrology*, 358(3), 221-239.
- 794 Kay, A.L., Davies, H.N., Bell, V.A., Jones, R.G., 2009. Comparison of uncertainty for climate change impacts:
795 flood frequency in England. *Clim. Change*. 92: 41-63. doi: 10.1007/s10584-008-9471-4.
- 796 Kay, A.L., Bell, V.A., Blyth, E. M., Crooks, S. M., Davies, H.N., Reynards, N. S. 2013. A hydrological perspective
797 on evaporation: historical trends and future projections in Britain. *Journal of Water and Climate Change*,
798 4(3) 193-208. doi: 10.2166/wcc.2013.014.
- 799 Kay, A.L., Jones, D.A., 2012. Transient changes in flood frequency and timing in Britain under potential
800 projections of climate change. *Int. J. Climatol.* 32: 489-502. doi: 10.1002/joc.2288.
- 801 Kay, A.L., Rudd, A.C., Davies, H.N., Kendon, E.J., Jones, R.G., 2015. Use of very high resolution climate model
802 data for hydrological modelling: baseline performance and future flood changes. *Clim. Change*. 133:193-
803 208. doi: 10.1007/s10584-015-1455-6.
- 804 Keller, V.D.J., Tanguy, M., Prosdocimi, I., Terry, J.A., Hitt, O., Cole, S.J., Fry, M., Morris, D.G., Dixon, H., 2015.
805 CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and
806 other applications. *Earth System Science Data*, 7(1), 143-155.
- 807 Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of
808 potential evapotranspiration under climate change. *Geophysical Research Letters*, 36(20).
- 809 Kim, K.B., Kwon, H.H., Han, D., 2016. Precipitation ensembles conforming to natural variations derived from a
810 regional climate model using a new bias correction scheme. *Hydrology and Earth System
811 Sciences*, 20(5), 2019-2034
- 812 Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D.,
813 van Meijgaard, E., Nikulin, G., Schär, C., Teichutard, R., Warrach-Sagi, K., Wulfmeyer, V., 2014.
814 Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM
815 ensemble. *Geosci. Model Dev.* 7, 1297-1333. doi: 10.5194/gmd-7-1297-2014.
- 816 Lafon, T., Dadson, S., Buys, G., Prudhomme, C., 2013. Bias correction of daily precipitation simulated by a
817 regional climate model: a comparison of methods. *Int. J. Climatol.* 33:1367-1381. doi:10.1002/joc.3518.

- 818 Larsen, M.A., Thejll, P., Christensen, J.H., Refsgaard, J.C., Jensen, K.H., 2013. On the role of domain size and
819 resolution in the simulations with the HIRHAM region climate model. *Climate dynamics*, 40(11-12), 2903-
820 2918.
- 821 Lockart, N., Willgoose, G., Kuczera, G., Kiem, A.S., Chowdhury, A.K., Manage, N.P., Zhang, L. and Twomey, C.,
822 2016. Case study on the use of dynamically downscaled climate model data for assessing water security
823 in the Lower Hunter region of the eastern seaboard of Australia. *Journal of Southern Hemisphere Earth
824 Systems Science*, 66(2), pp.177-202.
- 825 Macadam I., Argüeso, D., Evans, J.P., Liu, D.L., Pitman, A.J., 2016. The effect of bias correction and climate
826 model resolution on wheat simulations forced with a regional climate model ensemble. *Int. J. Climatol.*
827 doi: 10.1002/joc.4653.
- 828 Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J. Climate*.
829 26, 2137-2143. doi: 10.1175/JCLI-D-12-00821.1.
- 830 Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I.,
831 Soares, P.M., Hall, A. and Mearns, L.O., 2017. Towards process-informed bias correction of climate
832 change simulations. *Nature Climate Change*, 7(11), 3418.
- 833 Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W.,
834 Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M.,
835 Thiele-Eich., 2010. Precipitation downscaling under climate change: recent developments to bridge the
836 gap between dynamical models and the end user. *Rev. Geophys.* 48, RG3003.
837 doi:10.1029/2009RG000314.
- 838 Maraun, D., Widmann, M., 2015. The representation of location by a regional climate model in complex terrain.
839 *Hydrol. Earth Syst. Sci.* 19, 3449-3456. doi: 10.5194/hess-19-3449-2015.
- 840 Maraun, D., & Widmann, M., 2018. *Statistical downscaling and bias correction for climate research*. Cambridge
841 University Press.
- 842 Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, E., Hertig, E., Wibig, J., Huth, R., Wilcke,
843 R.A.I., 2015. VALUE: A framework to validate downscaling approaches for climate change studies.
844 *Earth's Future*. 3, 1-14. doi: 10.1002/2014EF000259.

- 845 Maurer, E.P., Pierce, D.W., 2014. Bias correction can modify climate model simulated precipitation changes
846 without adverse effect on the ensemble mean. *Hydrol. Earth Syst. Sci.* 18, 915-925. doi: 10.5194/hess-
847 18-915-2014.
- 848 Mendoza, P.A., Mizukami, N., Ikeda, K., Clark, M.P., Gutmann, E.D., Arnold, J.R., Brekke, L.D. and
849 Rajagopalan, B., 2016. Effects of different regional climate model resolution and forcing scales on
850 projected hydrologic changes. *Journal of Hydrology*, 541, pp.1003-1019.
- 851 Meredith, E.P., Maraun, D., Semenov, V.A., Park, W., 2015. Evidence for added value of convection-permitting
852 models for studying changes in extreme precipitation. *Journal of Geophysical Research:*
853 *Atmospheres*, 120(24), 12500-12513.
- 854 Monteith, J.L., 1965, July. Evaporation and environment. In *Symp. Soc. Exp. Biol*, 19, No. 205-23, p. 4).
- 855 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I: a discussion of
856 principles. *J. Hydrol.* 10, 282-290.
- 857 Nguyen, H., Mehrotra, R., Sharma, A., 2017. Can the variability in precipitation simulations across GCMs be
858 reduced through sensible bias correction?. *Climate Dynamics*, 1-19.
- 859 Onöl, B., 2012. Effects of coastal topography on climate: High-resolution simulation with a regional climate
860 model. *Clim. Res.* 52: 159-174. doi: 10.3354/cr01077.
- 861 Osborn, T.J. and Hulme, M., 1997. Development of a relationship between station and grid-box rainfall
862 frequencies for climate model evaluation. *Journal of Climate*, 10(8), pp.1885-1908.
- 863 Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005. Which potential
864 evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient
865 potential evapotranspiration model for rainfall-runoff modelling. *J. Hydrol.* 303, 290-306. doi:
866 10.1016/j.jhydrol.2004.08.026.
- 867 Parana Manage, N., Lockart, N., Willgoose, G., Kuczera, G., Kiem, A.S., Chowdhury, A.F.M., Zhang, L. and
868 Twomey, C., 2016. Statistical testing of dynamically downscaled rainfall data for the Upper Hunter
869 region, New South Wales, Australia. *Journal of Southern Hemisphere Earth System Science*, 66(2), 203-
870 227.

- 871 Pasten Zapata, E., 2017. An assessment of the impact of climate change on the efficiency and feasibility of four
872 run-of-the-river hydropower schemes in UK catchments (Doctoral dissertation, University of Sheffield).
- 873 Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate
874 models over Europe. *Theor. Appl. Climatol.* 99, 187-192. doi: 10.1007/s00704-009-0134-9.
- 875 Prein, A.F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E.,
876 Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., Jacob, D., 2015. Precipitation in the
877 EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Clim. Dyn.* doi:
878 10.1007/s00382-015-2589-y.
- 879 Prudhomme, C. and Davies, H., 2009. Assessing uncertainties in climate change impact analyses on the river
880 flow regimes in the UK. Part 1: baseline climate. *Climatic Change*, 93(1-2), 177-195. doi: DOI
881 10.1007/s10584-008-9464-3
- 882 Prudhomme, C., Haxon, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., Kelvin, J., Mackay, J., Wang,
883 L., Young, A., Watts, G., 2013. Future flows hydrology: an ensemble of daily river flow and monthly
884 groundwater levels for use for climate change impact assessment across Great Britain. *Earth Syst. Sci.*
885 *Data* 5: 101-107. doi:10-5194/essd-5-101-2013.
- 886 Qian, B., Wang, H., He, Y., Liu, J., De Jong, R., 2015. Projecting spring wheat yield changes on the Canadian
887 Prairies: effects of resolutions of a regional climate model and statistical processing. *Int. J. Climatol.* doi:
888 10.1002/joc.4571.
- 889 Robinson, E., Blyth, E., Clark, D., Comyn-Platt, E., Finch, J., Rudd, A., 2017a Climate hydrology and ecology
890 research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2. NERC
891 Environmental Information Data Centre 7 February 2017, [https://doi.org/10.5285/b745e7b1-626c-4ccc-](https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900)
892 [ac27-56582e77b900](https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900)
- 893 Robinson, E.L., Blyth, E.M., Clark, D.B., Finch, J., Rudd, A.C., 2017b. Trends in atmospheric evaporative
894 demand in Great Britain using high-resolution meteorological data. *Hydrology and Earth System*
895 *Sciences*,21(2), 1189.
- 896 Rockel, B., 2015. The regional downscaling approach: a brief history and recent advances. *Curr. Clim. Change*
897 *Rep.* 1:22-29. doi: 10.1007/s40641-014-0001-3.

- 898 Rockel, B., Will, A., Hense, A., 2008. Special issue on regional climate modelling with COSMO-CLM (CCLM).
899 Meteorol. Z. 17, 477-485.
- 900 Rummukainen, M. B., 2016. Added value in regional climate modeling. WIREs Clim. Change 7:145-159. doi:
901 10.1002/wcc.378
- 902 Rummukainen, M.B., Rockel, L., Bärring, J., Christensen, J.H., Reckermann, M., 2015. 21st century challenges
903 in Regional Climate Modeling. Bull. Amer. Meteor. Soc. doi:10.1175/BAMS-D-14-00214.1.
- 904 Samuelsson, P., Jones, C.G., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E.,
905 Nikulin, G., Wyser, K., 2011. The Rossby Centre Regional Climate Model RCA3: model description and
906 performance. Tellus 63A:4-23. doi: 10.1111/j.1600-0870.2010.00478.x.
- 907 Seiller, G. and Anctil, F., 2016. How do potential evapotranspiration formulae influence hydrological projections?.
908 Hydrological Sciences Journal, 61(12), 2249-2266.
- 909 Scharffenberg, W.A., 2013. Hydrologic Modeling System HEC-HMS User's Manual Version 4.0. Hydrologic
910 Engineering Center, US Army Corps of Engineers; 442pp. Davis, USA.
- 911 Skamarock, W.C., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Powers, J., 2008. A description of the
912 advanced research wrf version 3, Tech. Rep., NCAR Technical Note 475, 113pp.
- 913 Tanguy, M., Dixon, H., Prosdociimi, I., Morris, D. G., Keller, V. D. J. (2014). Gridded estimates of daily and
914 monthly areal rainfall for the United Kingdom (1890-2012) [CEH-GEAR]. NERC Environmental
915 Information Data Centre. <https://doi.org/10.5285/5dc179dc-f692-49ba-9326-a6893a503f6e>
- 916 Teng, J., Potter, N.J., Chiew, F.H.S., Zhang, L., Wang, B., Vaze, J., Evans, J.P., 2015. How does bias correction
917 of regional climate model precipitation affect modeled runoff? Hydrol. Earth Syst. Sci. 19, 711-728. doi:
918 10.5194/hess-19-711-2015.
- 919 Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-
920 change impact studies: Review and evaluation of different methods. J. Hydrol. 456, 12-29. doi:
921 10.1016/j.jhydrol.2012.05.052.
- 922 Tolika, K., Anagnostopoulou, C., Velikou, K. and Vagenas, C., 2016. A comparison of the updated very high
923 resolution model RegCM3_10km with the previous version RegCM3_25km over the complex terrain of
924 Greece: present and future projections. Theoretical and applied climatology, 126(3-4), pp.715-726.

- 925 Tölle, M.H., Schefczyk, L. and Gutjahr, O., 2017. Scale dependency of regional climate modeling of current and
926 future climate extremes in Germany. *Theoretical and Applied Climatology*, pp.1-20.
- 927 Torma, C., Giorgi, F., Coppola, E., 2015. Added value of regional climate modeling over areas characterized by
928 complex terrain – Precipitation over the Alps. *J. Geophys. Res. Atmos.* 120:3957-3972. doi:
929 10.1002/2014JD022781.
- 930 van Meijgaard, E., van Ulft, L.H., Lenderink, G., de Roode, S.R., Wipfler, L., Boers, R., Timmermans, R.M.A.,
931 2012. Refinement and application of a regional atmospheric model for climate scenario calculations of
932 Western Europe. *Climate changes Spatial Planning publication: KvR 054/12*.
- 933 Velázquez, J.A., Troin, M., Caya, D., Brissette, F., 2015. Evaluating the time-invariance hypothesis of climate
934 model bias correction: implications for hydrological impact studies. *J. Hydrometeor.* 16, 2013-2026. doi:
935 10.1175/JHM-D-14-0159-1.
- 936 Walsh, C. L., Blenkinsop, S., Fowler, H. J., Burton, A., Dawson, R. J., Glenis, V., Manninga, L.J. and Kilsby, C.
937 G., 2015. Adaptation of water resource systems to an uncertain future. *Hydrology & Earth System
938 Sciences Discussions*, 12(9).
- 939 Wang, H. M., Chen, J., Xu, C. Y., Chen, H., Guo, S., Xie, P., & Li, X., 2019. Does the weighting of climate
940 simulations result in a better quantification of hydrological impacts?. *Hydrology and Earth System
941 Sciences*, 23(10), 4033-4050.
- 942 Wetterhall, F., Pappenberger, F., He, Y., Freer, J. and Cloke, H.L., 2012. Conditioning model output statistics of
943 regional climate model precipitation on circulation patterns. *Nonlinear Processes in Geophysics*, 19(6),
944 pp.623-633.
- 945 Wilby, R. L., and Harris, I., 2006. A framework for assessing uncertainties in climate change impacts: Low-flow
946 scenarios for the River Thames, UK. *Water Resources Research*, 42(2).
- 947 Wilcke, R.A.I., Mendlik, T., Gobiet, A., 2013. Multi-variable error correction of regional climate models. *Climatic
948 Change*. 120, 871-887. doi: 10.1007/s10584-013-0845-x.
- 949 Yang, W., Andréasson, J., Graham, L.P., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling
950 to improve usability of regional climate model projections for hydrological climate change impacts
951 studies. *Hydrology Research*, 41(3-4), 211-229.

952 Zittis, G., Bruggeman, A., Camera, C., Hadjinicolaou, P., Lelieveld, J., 2017. The added value of convection
 953 permitting simulations of extreme precipitation events over the eastern Mediterranean. *Atmospheric*
 954 *Research*, 191, 20-33.

955

956 Figure 1. Location of the study catchments (outlet marked with a triangle) and the RCM grid boxes used for their
 957 simulation. The 0.11° and 0.44° grid boxes are shown with solid and dashed lines, respectively

958

959 Figure 2. Results of the temperature performance measures, described on Table 3, for the a) upper Thames, b) Glaslyn, c)
 960 Calder and d) Coquet catchments. Filled symbols represent the 0.11° RCMs and empty symbols represent the 0.44° RCMs.
 961 Please note the differences in the y-axis (BC = Bias corrected)

962

963 Figure 3. Temperature percentile biases for the uncorrected and bias-corrected RCMs. The solid fill represents the spread
 964 form the 0.44° RCMs and the dotted fill is the spread from the 0.11° RCMs

965

966 Figure 4. Results of the precipitation performance measures for the upper Thames catchment. Filled symbols represent
 967 the 0.11° RCMs and empty symbols represent the 0.44° RCMs. Please note the differences in the y-axis. For definitions of
 968 the performance measures refer to Table 3 (BC-1G = Bias corrected using the Gamma distribution QM approach, BC-2G =
 969 Bias corrected using the Double Gamma distribution approach)

970 Figure 5. Precipitation percentile biases for the uncorrected and bias-corrected RCMs using the Gamma distribution
 971 (GQM) and Double Gamma distribution (DGQM) QM. The solid fill represents the spread of the 0.44° RCMs and the dotted
 972 fill the spread of the 0.11° RCMs. The 90th precipitation percentile is represented by a vertical dotted line

973

974 Figure 6. Results of the river flow performance measures for the upper Thames catchment. Filled symbols represent the
 975 0.11° RCMs and empty symbols represent the 0.44° RCMs. Please note the differences in the y-axis. For definitions of the
 976 performance measures refer to Table 3 (GQM = Gamma distribution Quantile Mapping and DGQM = double Gamma
 977 distribution Quantile Mapping)

978

979 Figure 7. Flow duration curve biases from using the uncorrected and bias-corrected temperature and precipitation
 980 simulations. The 0.44° RCMs spread is shown with a solid fill, the 0.11° RCMs spread with a dotted fill and the reference
 981 FDC with a solid line. (GQM = Gamma distribution Quantile Mapping and DGQM = double Gamma distribution Quantile
 982 Mapping)

983

984 Figure S1. Similar to Figure 5 but for the Glaslyn catchment

985

986 Figure S2. Similar to Figure 5 but for the Calder catchment

987

988 Figure S3. Similar to Figure 5 but for the Coquet catchment

989

990 Figure S4. Similar to Figure 7 but for the Glaslyn catchment

991

992 Figure S5. Similar to Figure 7 but for the Calder catchment

993

994 Figure S6. Similar to Figure 7 but for the Coquet catchment

995

996

997

998

999

- 0.11° and 0.44° RCMs are compared in four catchments using climate and flow indices

1000

- The 0.11° simulations had superior skill in one catchment with complex topography

1001

- The RCM flow simulation range is large for all catchments at both resolutions

1002

- Bias correction improves the monthly but not the daily temporal variability

1003

- Double Gamma quantile mapping outperforms the single Gamma quantile mapping

1004

1005

1006 Declaration of interests

1007

1008

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1009

1010

1011

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

1012

1013

1014

1015

1016

1017

1018

1019 Author statement

- 1020 Pasten-Zapata: Funding acquisition, conceptualization, methodology, investigation, formal analysis and writing
1021 original-draft
- 1022 Jones: Conceptualization, resources, writing original-draft, writing review and editing, supervision
- 1023 Moggridge: Conceptualization, resources, writing original-draft, writing review and editing, supervision
- 1024 Widmann: Conceptualization, writing original-draft, writing review and editing
- 1025
- 1026

Journal Pre-proofs